SOME LEAST-SQUARES TRANSFORMATIONS OF REGRESSION ESTIMATORS OF ORTHOGONAL FACTORS

Emil F. Heermann
U. S. Army Personnel Research Office

This paper provides a brief introduction to the factor analysis model used by psychologists and presents some methods for transforming regression estimators of orthogonal factors to a more useful form.

I. THE FACTOR ANALYSIS MODEL. Factor analysis is a mathematical method for representing n correlated measures obtained on N individuals by means of a linear combination of hypothetical measures, called factors. Psychologists draw a distinction between common factors, which are related to two or more measures, and specific factors which are related to only one of the original measures. In general, then, the linear model postulates n specific factors and r common factors where r is less than the number of measures n. The sampling theory for such a model is not yet thoroughly understood and this paper will confine itself to algebraic rather than statistical considerations.

The linear model of factor analysis can be most conveniently presented in the form of matrix equations:

(1) $$Z = Z_c F' + Z_s S' \text{ , where}$$

Z = N x n matrix of observed measures (zero mean, unit variances)

Z_c = N x r matrix of common factor scores

F' = r x N matrix of common factor weights

Z_s = N x n matrix of specific factor scores

S' = n x n diagonal matrix of specific factor weights

Terms on the right side of the equation are all unknown. In (1) it is assumed that:

(2) $$\frac{1}{N} Z'_c Z_c = I_r$$

(3) $$\frac{1}{N} Z'_s Z_s = I_n$$

(4) $$\frac{1}{N} Z'_c Z_s = 0$$

These assumptions allow us to find F' and S' from the observed measures, but before we examine this procedure it is important to indicate a rather serious difficulty with our linear model. When we postulate $n + r$ factors to explain n measures this invariably means that there exist multiple solutions for an individual's factor scores. Solutions exist because the linear equations are consistent but the rank of the coefficient matrix does not allow unique solutions. To get around this difficulty psychologists have utilized least-squares procedures for estimating the factor scores from the observed measures, and we shall examine the defects of these estimators at a later point.

We now return to the problem of finding F' and S' from the observed data. Because of (2), (3) and (4) we may write:

$$R = \frac{1}{N} Z'Z = FF' + SS', \text{ or}$$

$$R - SS' = FF'$$

Since SS' is a diagonal matrix, we begin by estimating values for SS' and then obtain the eigenvectors and eigenvalues of

$$R - SS' = A\,K\,A', \text{ where}$$

$A = n \times r$ matrix of eigenvectors

$K = r \times r$ matrix of eigenvalues

Now, $F = AK^{1/2}$, and the sums of squares of the rows of F are used to obtain new values of SS', and the process is continued until $R - SS'$ is fit with the minimum rank F. This whole process may be thought of as finding a set of values for SS', such that $R - SS'$ is of minimum rank.

It will be noted that F is arbitrary up to an orthogonal transformation and it is necessary to postulate some method for finding a psychologically meaningful F. Psychologists follow L. L. Thurstone here and attempt to transform the arbitrary F, so that it approximates "simple structure", i.e., has a maximal number of zero or near-zero entries. To accomplish this we find some orthogonal transformation, λ, such that

$$F\lambda = F_R$$

where $\lambda'\lambda = I$, and

F_r approximates "simple structure". A considerable number of computer programs exist which determine by analytical means the best transformation λ.

II. <u>ESTIMATION OF FACTOR SCORES.</u> Assuming now that F_r is adequately fixed, our problem is then to find the values of Z_c and Z_s that will satisfy the linear model specified in (1). Since an infinite set of such factor scores exist, psychologists commonly turn to a regression method for estimating Z_c for a fixed Z_s. We are primarily interested here in finding values for Z_c, and the least-squares solution for Z_c is easily found to be

$$ZR^{-1}F_R = Z_c .$$

But unlike the factor scores Z_c, these estimators are intercorrelated, because

(5) $$\frac{1}{N}\hat{Z}'_c\hat{Z}_c = F'_R R^{-1} F_R,$$ the

covariance matrix of the least-squares estimator is not diagonal. Adjusting the covariance matrix so that we have an intercorrelation matrix gives:

$$R_L = D_e^{-1} F'_R R^{-1} F_R D_e^{-1},$$

where

D_e = r x r diagonal matrix, formed from the square roots of diagonal entries of (5), the covariance matrix.

A further practical difficulty is that regression estimators are not univocal, i.e., each estimator is correlated with more than one common factor. This can be seen by examining the matrix of correlations between the factor scores and the least-squares estimators, which is

$$R_{fL} = D_e R_L.$$

It is not too difficult to derive transformations of the beta weights used in the regression estimators that remove these defects, but it is not possible by means of a single transformation to simultaneously remove both defects.

To adjust for non-orthogonality, we first find an arbitrary set of orthogonal vectors that serves as a vector basis for the regression estimators. To do this we find an $r \times r$ T, such that

$$R_L = TT'$$

The elements of T represent the correlations of the least-squares estimators with the r orthogonal axes. The correlations of the factors with the orthogonal axes are given by

$$R_{fT} = D_e T,$$

and to find the best set of orthogonal axes we must transform T so that $D_e T$ most closely approximates the diagonal matrix D_e. In matrix algebra, we wish to find a λ, such that if

$$R_{fT}\lambda - D_e = E, \text{ then}$$

$$\text{trace } E'E = \text{minimum, and}$$

$$\lambda' \lambda = I.$$

From a theorem proved by B. F. Green (1) we find

$$\Lambda = (T' D_e^4 T)^{-1/2} T' D_e^2 \quad .$$

Now if the requirements of the psychologist are such that it is more important to have univocal rather than orthogonal estimators than a somewhat different procedure is required. In this case we determine so that if

$$R_{fT} \Lambda - D_e = E \, ,$$

$$\text{trace } E'E = \text{minimum, and}$$

$$\Lambda'_{\cdot i} \Lambda_{\cdot i} = 1, \ (i = 1, 2, \ldots r), \text{ where}$$

$$\Lambda_{\cdot i} = i^{th} \text{ column of } \Lambda .$$

As can be seen the orthogonality restraint on Λ has been relaxed. The normal equations for these conditions are

(8) $$(T' D_e^2 T - \gamma) \Lambda = T' D_e^2 \quad ,$$

where γ = r x r matrix of Lagrangian Multipliers. The above equation does not allow us to find a matrix expression for Λ which does not also involve γ . Hence we pursue a simpler approach which involves a least-squares solution for Λ followed by imposing the restraints that each column of Λ have sums of squares equal to unity. With this approach we find that

$$\Lambda_n = T^{-1} D, \text{ where}$$

D is the diagonal matrix of constants needed to adjust T^{-1} so that the column sums of squares equal unity. When the least-squares estimators are transformed by Λ_N we find that the matrix of correlations between Z_c and our transformed estimators is $R_{fN} = D_e D$, which is quite obviously a diagonal matrix.

III. AN APPLICATION OF THE TRANSFORMATIONS. In 1954 APRO carried out a factor analysis of visual acuity tests administered during dark adaption. The examinees in this experiment were 100 soldiers from Fort Myer, Virginia. The design of the experiment called for pre-adapting subjects to a high brightness level and then during the period of dark adaption the examinees were tested on visual acuity targets presented at scotopic, mesopic, and low photopic brightness. The various acuity tests (Modified Landolt Ring and Chevron Contrast Adoption tests) were presented in a modified Armed Forces Vision Test.

The 35 variable intercorrelation matrix obtained in this experiment yielded 8 orthogonal factors which were rotated to orthogonal simple structure. To illustrate the derivations presented in this paper, a submatrix was selected from the 35 x 8 complete factor matrix. Table 1 gives this submatrix.

Table 1

Illustrative Factor Matrix

Variable Name		I	II	III	IV
Landolt Scotopic	1	.84	.07	.13	.34
Landolt Low Photopic	2	.16	.85	.03	.13
Chevron Scotopic	3	.28	.32	.12	.75
Chevron Mesopic	4	.01	.03	.79	.29

Factor I was interpreted to represent "Rod-Adapted Resolution"; Factor II was interpreted to represent "Cone Adapted Resolution"; Factor III was interpreted to represent "Cone Adapted Brightness Discrimination", and Factor IV was interpreted as "Rod-Adapted Brightness Discrimination".

Table 2 compares the classic regression equations approach to estimating these four factors with the methods derived in this paper. As noted previously the least-squares estimators are intercorrelated and fail to be univocal. The univocal estimators achieve an ideal pattern of correlations with the factors but they are intercorrelated somewhat more than the least-squares estimators. The orthogonal estimators, while uncorrelated, are not univocal although they are closer to being univocal than the least-squares estimators. Inevitably then, choice of any one solution means that certain defects must be tolerated. From the results presented in Table 2 it would appear that the orthogonal estimators represent something of a compromise between the maximum validity of the least-squares estimators and purity of the univocal estimators.

(1) Green, B. F. The orthogonal approximation of an oblique structure in factor analysis. Psychometrika, 1952, 17 429-440.